Executive Summary

Writing Assessment at Westminster College

October 15, 2012

Theresa Adams, Ph.D., Assistant Professor and Director, Writing Across the Curriculum
Ray Brown, Ph.D., Director of Institutional Research

Paired writing samples were reviewed by a committee of faculty members during the spring semester of the 2011-2012 academic year.   The writing assessment process included several new features this year.   First, the assessment was performed on Assessment Day, rather than on a weekend or during a class break. Second, the English Department faculty served as a core group for the assessment process supplemented by full-time teaching faculty colleagues from several other departments.   Third, the assessment was performed using a newly formulated rubric created by the Director of Writing Across the Curriculum in consultation with the Assessment Committee.

Each of the writing samples included two papers, one written by the student during their first semester on campus in Freshman Seminar class and the other written while enrolled in an upper level Tier course.   Committee members rated the writing samples in three categories [Thesis and Support, Organization, and Style and Grammar/Mechanics/Usage] using a five-point scale of 1 = weak, 3 = adequate, and 5 = strong.

The writing samples were tagged by the entry year when the student was a first time freshman and this entry year served as one variable in the analysis.   The faculty ratings also served as variables in the analysis and Table #1 provides the mean ratings for papers.   Ratings were then entered into Excel by a student assistant and imported into SPSS PASW version 18.0 for statistical analysis.
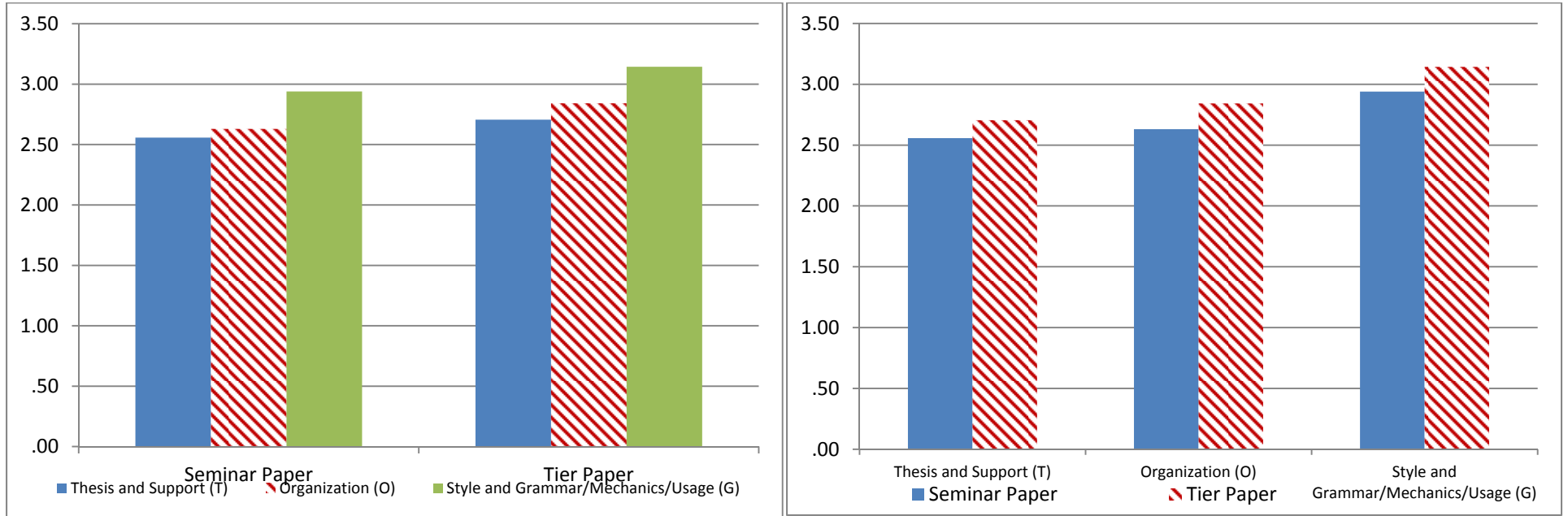
Analysis for Seminar and Tier Papers

Mean values were compared by the origin of the papers in either the Freshman Seminar or the upper level Tier course [Table #1a].   Mean ratings were lowest for the "Thesis and Support" category and highest for the "Style and Grammar/Mechanics/Usage" category for both groups of papers [see also, Chart #1a]. Also, the means for all three categories were higher for the papers collected from the Tier classes than for those collected from the Freshman Seminar classes [see also, Chart #1b].

Table #1a: Means by Origin of Writing Sample

		Thesis and Support (T)	Organization (O)	Style and Grammar/Mechanics/Usage (G)
Seminar Paper	Means	2.56	2.63	2.94
Tier Paper	Means	2.71	2.84	3.14
	Difference in Means	**.15**	**.21**	**.20**

Charts #1a and #1b

Analysis of variance was employed to test for significant differences in means for any of the three rating categories between the Seminar and Tier papers.   The results are displayed in Table #2.   None of the resulting F-values were found to be significantly different at the α = 0.05 level of significance.

ANOVA Table #2

			Sum of Squares	df	Mean Square	F	Sig.
Thesis and Support (T) * Seminar or Tier Writing Sample	Between Groups	(Combined)	1.672	1	1.672	1.456	.229
	Within Groups		344.530	300	1.148		
	Total		346.202	301			
Organization (O) * Seminar or Tier Writing Sample	Between Groups	(Combined)	3.401	1	3.401	3.010	.084
	Within Groups		338.933	300	1.130		
	Total		342.334	301			
Style and Grammar/Mechanics/Usage (G) * Seminar or Tier Writing Sample	Between Groups	(Combined)	3.147	1	3.147	3.333	.069
	Within Groups		283.293	300	.944		
	Total		286.440	301			

Analysis by Year of Entry to Westminster College

Mean values were compared by the year in which students entered Westminster College as first time freshmen [Tables #3a and #3b].
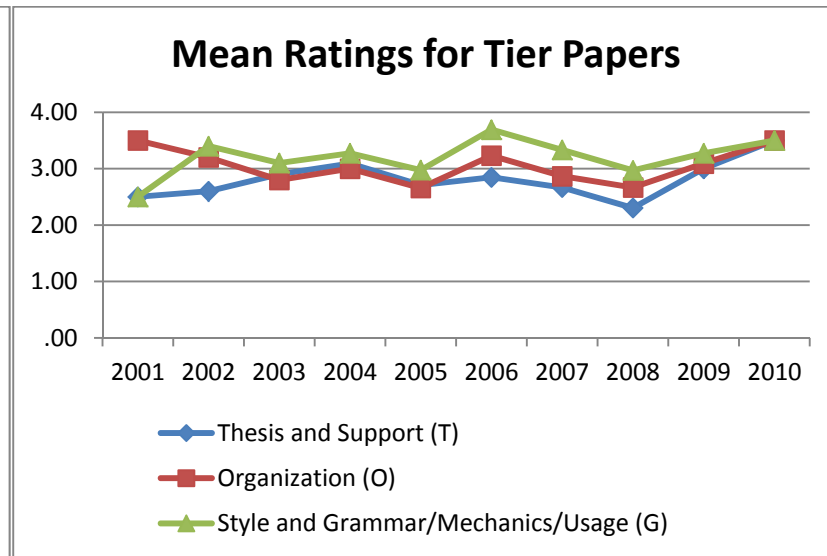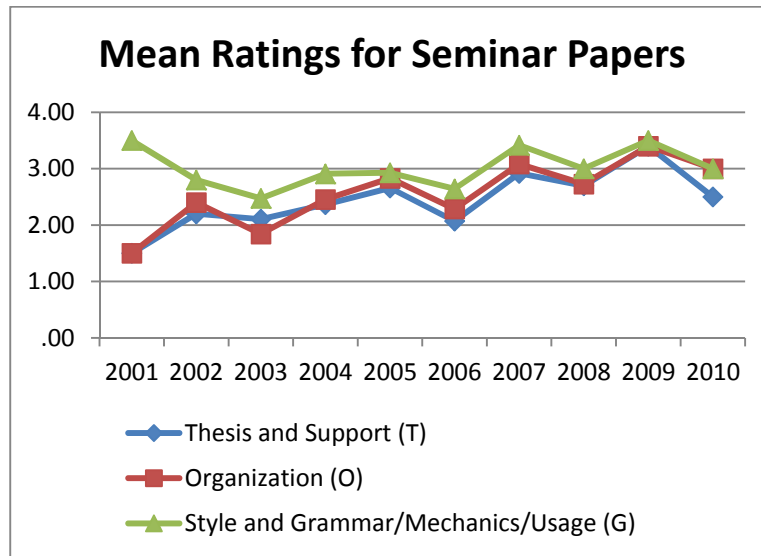
Table #3a: Mean Ratings for Seminar Papers

Entry Year	n's	Thesis and Support (T)	Organization (O)	Style and Grammar/Mechanics/Usage (G)
2001	2	1.50	1.50	3.50
2002	5	2.20	2.40	2.80
2003	19	2.11	1.84	2.47
2004	11	2.36	2.45	2.91
2005	41	2.66	2.83	2.93
2006	14	2.07	2.29	2.64
2007	12	2.92	3.08	3.42
2008	33	2.70	2.73	3.00
2009	10	3.40	3.40	3.50
2010	2	2.50	3.00	3.00

Table #3b: Mean Ratings for Tier Papers

Entry Year	n's	Thesis and Support (T)	Organization (O)	Style and Grammar/Mechanics/Usage (G)
2001	2	2.50	3.50	2.50
2002	5	2.60	3.20	3.40
2003	20	2.90	2.80	3.10
2004	11	3.09	3.00	3.27
2005	41	2.71	2.66	2.98
2006	13	2.85	3.23	3.69
2007	15	2.67	2.87	3.33
2008	33	2.30	2.67	2.97
2009	11	3.00	3.09	3.27
2010	2	3.50	3.50	3.50

Charts #2a and #2b

Mean Ratings for Seminar Papers

Mean Ratings for Tier Papers

Thesis and Support (T)
Organization (O)
Style and Grammar/Mechanics/Usage (G)

Visually the general pattern is one of stability in means over time, though there does appear to be an obvious upward trend in for "Organization" and "Thesis and Support" ratings given to the Seminar papers.   Caution is certainly warranted since the number of writing samples for some years is quite small.

Analysis of variance was also employed to test for significant differences in means by entry year for any of the three rating categories.   These results are displayed in Table #3.   None of the resulting F-values were found to be significantly different at the $\alpha = 0.05$ level of significance.

ANOVA Table #3

			Sum of Squares	df	Mean Square	F	Sig.
Thesis and Support (T) * Entry Year	Between Groups	(Combined)	12.867	9	1.430	1.252	.263
	Within Groups		333.335	292	1.142		
	Total		346.202	301			
Organization (O) * Entry Year	Between Groups	(Combined)	14.435	9	1.604	1.428	.175
	Within Groups		327.899	292	1.123		
	Total		342.334	301			
Style and Grammar/Mechanics/Usage (G) * Entry Year	Between Groups	(Combined)	9.167	9	1.019	1.073	.383
	Within Groups		277.273	292	.950		
	Total		286.440	301			

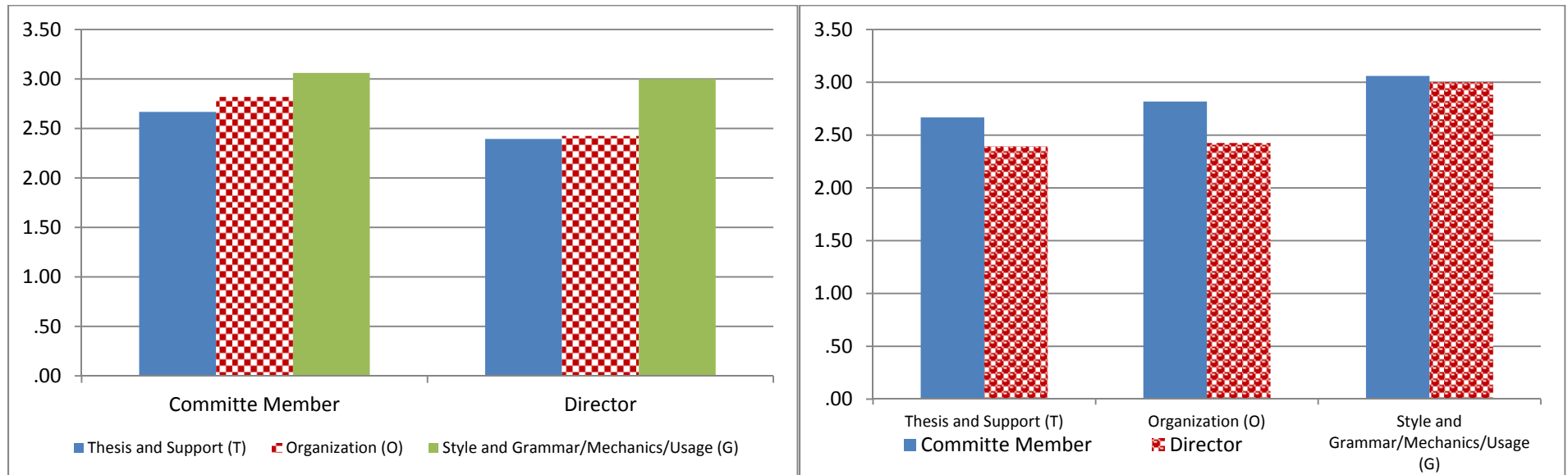**Analysis of Ratings for Faculty and the Director of Writing Assessment**

Thirty-three papers were pulled from the larger pool and also received ratings by the Director of Writing Assessment.   This step was completed to provide an additional set of ratings that could be compared with those given by the faculty committee.   Table #4 and corresponding Charts #3a and #3b provide a summary of the mean ratings.

Ratings by the Director were lower with the exception of those awarded for Style and Grammar/Mechanics/Usage for Tier papers.   In this instance the mean of ratings by the Director was 3.13 and the corresponding mean rating by the Committee Members was 2.94.   In all other instances, the mean of Committee Members was higher and the mean differences ranges from - 0.13 for the Thesis and Support on the Tier papers to – 0.53 for the Organization category on Seminar papers.

Table #4a

			Thesis and Support (T)	Organization (O)	Style and Grammar/Mechanics/Usage (G)
Seminar Paper	Committee Members	Means	2.71	3.00	3.18
	Director	Means	2.29	2.47	2.88
		Difference	**-.41**	**-.53**	**-.29**
Tier Paper	Committee Members	Means	2.63	2.63	2.94
	Director	Means	2.50	2.38	3.13
		Difference	**-.13**	**-.25**	**.19**

Charts #3a & #3b

Once again, analysis of variance was employed to test for significant differences in means for faculty ratings and those of the Director of Writing Assessment. These results are displayed in Table #6. None of the resulting F-values were found to be significantly different at the α = 0.05 level of significance.

ANOVA Table #6

			Sum of Squares	df	Mean Square	F	Sig.
Thesis and Support (T) * Seminar or Tier Writing Sample	Between Groups	(Combined)	.064	1	.064	.066	.798
	Within Groups		62.375	64	.975		
	Total		62.439	65			
Organization (O) * Seminar or Tier Writing Sample	Between Groups	(Combined)	.913	1	.913	1.154	.287
	Within Groups		50.618	64	.791		
	Total		51.530	65			
Style and Grammar/Mechanics/Usage (G) * Seminar or Tier Writing Sample	Between Groups	(Combined)	.000	1	.000	.000	.994
	Within Groups		55.939	64	.874		
	Total		55.939	65			

Summary and Conclusions

The writing assessment process was modified in AY2011. Informal feedback from participating faculty members and discussion between the Director of Writing Across the Curriculum and the Assessment Committee was positive. It appears that the three major changes were successful and should be continued. It was easier to find faculty colleagues willing to participate on Assessment Day. It was also helpful to have the English Department faculty participating, since this contributed to a shared vocabulary for talking about writing and it was easier to discuss papers during the norming session. Finally, the more streamlined rubric, was much more functional than the previous version since faculty felt it was more aligned with "intuitive" grading.

The following conclusions can also be drawn from the analysis:
- The quality of writing for Westminster College students does improve between their Seminar and Tier classes, though not significantly.
- The quality of writing for successive entering classes appears to be relatively stable over time.
- Faculty ratings are not significantly different from those given on the same writing samples by the director of writing assessment.

Appendix

Faculty ratings of student writing samples from the spring 2012 writing assessment was also used in a second study to see what relationships might exist with scores on the Collegiate Learning Assessment (CLA).   Ratings for writing samples and CLA scores were matched for fifty-eight students. Thirty-eight (38) of the students completed CLA performance tasks and twenty students (20) analytic writing tasks.   Mean values for the variables are replicated in Table #7.

Table #7   Descriptive Statistics

		Mean	Std. Deviation	N
Ratings of Writing Samples by Westminster Faculty	Thesis and Support (T)	2.74	1.069	58
	Organization (O)	2.97	.973	58
	Style and Grammar/Mechanics/Usage (G)	3.09	.978	58
CLA sub-scores	Analytic Reasoning PT	3.68	.989	38
	Writing Effectiveness PT	3.89	.924	38
	Writing Mechanics PT	4.00	.986	38
	Problem Solving PT	3.58	1.004	38
	Analytic Reasoning MA	3.70	.657	20
	Writing Effectiveness MA	3.70	.657	20
	Writing Mechanics MA	4.00	.795	20
	Analytic Reasoning CA	3.10	.553	20
	Writing Effectiveness CA	3.30	.470	20
	Writing Mechanics CA	3.70	.657	20
	Time on Task for CLA in minutes	45.93	20.320	58
	ACT Composite Scores	27.31	3.213	58

The non-parametric Kendall rank correlation coefficient in SPSS PASW v. 18 was used to measure the association between variables.   Specifically, the coefficients calculated are measures of the similarity of the orderings of the data when ranked by each of the quantities.   The Kendall test is an alternative to Spearman's rank correlation coefficient in situations where there are a number of ties in ranks assigned to the variables.

Significant values for pairs of variables are identified in Table #8.   The calculated values are greater than what would be expected by chance alone.   However, caution must be exercised since the sample size is so small [58 total students, 38 with CLA performance task sub scores and 20 students with CLA analytic writing task sub scores].   While statistical significance does not necessarily imply that results are substantive or important and there is no reason to assume a causal link, it is none the less interesting that there is an apparent positive relationship between student sub scores on the CLA performance tasks and the faculty ratings of their writing ability.   Further, it is also interesting that there are no significant correlations between sub scores on CLA analytic writing tasks and the faculty ratings.   Continued monitoring of results and performance of similar analysis in the future is certainly warranted.

Table #8   Significant Correlations

	Ratings of Student Writing by WC Faculty		
Ratings of Student Writing by CLA	Thesis and Support (T)	Organization (O)	Style and Grammar/Mechanics/Usage (G)
Analytic Reasoning PT	*		
Writing Effectiveness PT	**	**	
Writing Mechanics PT	*	*	
Problem Solving PT	*		
Analytic Reasoning MA			
Writing Effectiveness MA			
Writing Mechanics MA			
Analytic Reasoning CA			
Writing Effectiveness CA			
Writing Mechanics CA			
Time on Task	*		
ACT			

** Correlation is significant at the 0.01 level (2-tailed).
*   Correlation is significant at the 0.05 level (2-tailed).